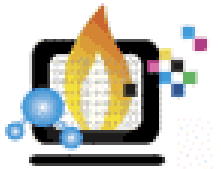

Security Vision and Use Cases

David Leahy, Sandra Bittner, Gregor von Laszewski, Karen Schuchardt

djleahy@sandia.gov, [bittner, gregor]@mcs.anl.gov,
karen.schuchardt@pnl.gov

Sandia, Argonne, and Pacific Northwest National Laboratories

June 13, 2002

Security Goals

- Security First
  - The CMCS project must consider security issues in its design from the very beginning. Too often a complicated and ambitious project can find out late in the development cycle that neglected security concerns wind up requiring huge resource allocations or significant de-featuring of the project.
- Low Barrier to Entry
  - CMCS must be easy to use in order to be successful with the larger goal of attracting researchers and information science developers. Analysis of the security challenges associated with possible avenues of development have helped to define the scope and the timeline of the CMCS project as a whole.
- Facilitate Research, Rather tthan Obstruct It
  - Our goal is a security implementation that facilitates research by assuring the users that their concerns are addressed, while minimizing the costs of security development and administration.

Security Landscape and Vision

- (A) Data-Centric Project
  - While the CMCS has many facets, at its core it is a data-centric content management and publishing system. Thus, we have initially focused our security effort on securing a WebDAV server.
- (B) Secured Workspaces
  - CMCS will supply chemical sciences research projects with secured workspaces and secured communications. Initially, widely-adopted commodity protocols such as HTTP and SSL will be used for prototyping.
- (C) Data Publishing with Non-Repudiation
  - In the first prototype, published data will be globally readable. It will be digitally signed and time-stamped, and safely archived. This comprises a high standard for permanent data storage and publishing.
- (D) Integration of Remote Computing Resources
  - CMCS will follow a Web Services model, where remote sites that access the CMCS data store will typically be responsible for securing their own assets; CMCS will secure their data.
- (E) Notification
  - CMCS will provide notifications on a subscription basis to users and applications that want to know about updates. Resulting security implications (especially privacy concerns) will be addressed.

Typical Use Case

- Project Scales
  - We envision typical collaborations consisting of projects with a few to a dozen researchers sharing a common, password-secured workspace. The project coordinator will administrate access through traditional Access Control Lists.
- Data Scales
  - Typical databases of chemical information will rarely exceed a few MBs. Some of our users (e.g., Direct Numerical Simulation researchers) own very large data sets; in these cases we anticipate that CMCS will host much smaller quantities of metadata, while the primary datasets will be stored outside of the CMCS. For the prototype, the size of the secured data store will be manageable for a single DAV server.
- Usage Patterns
  - Work-in-progress will be shared privately between project members, with authorization for outsiders being granted only at the discretion of the project coordinator. For some users, this will be the primary mode of participation in CMCS. Other users will use their secured workspaces as a stepping stone to permanently publishing their digital content on CMCS. Published data will have to meet requirements defined in the CMCS Publication Policy, and will always be globally accessible.

Active Tables Use Case

- Project Scales
  - Active Tables is a more ambitious CMCS Use Case. It will be a robust (many simultaneous users) stand-alone web service and will be publicly accessible. Active Tables could scale up to thousands of participants. Access to the Active Tables web service is outside of the domain of the CMCS; however, CMCS provides a data storage interface to the Active Tables web service. This will allow Active Table users to store and share their results using the CMCS infrastructure.
- Data Scales
  - The core of thermodynamic data used by Active Tables will be of the order of MBs. However, collectively the workspaces of users could amount to a considerably larger amount of dynamic data. Thus, Active Table's participation in CMCS will provide a robust test case for performance and reliability of the CMCS data storage interface.
- Usage Patterns
  - Work-in-progress will typically be held privately by individual Active Table users. Some sophisticated users will take advantage of CMCS as a publishing forum, permanently publishing their calculation results on CMCS.

Notification Use Case

- Feature Description
  - CMCS notification services will allow users to be informed when data repositories of interest are updated or modified in any way. Users can subscribe for notifications from both published and unpublished repositories.
- Security Implications
  - The Notification server will need to ensure that only those with access to specific data repositories may receive notification about updates to that information. Care must be taken not to open security holes by circumventing the security of the Data Storage Interface itself.
- Configuration Implications
  - The Notification server will generally be privy to information that is not to be made publicly available. As such, it will be placed and administrated in a secure environment such as that of the WebDAV data repository itself.

Future Plans for CMCS Security

- Implementation-Independent Security Model (IISM)
  - We plan to abstract the authentication and authorization process from the specific technologies that need to be secured (initially, the WebDAV data repository, and the CMCS portal). This abstraction will facilitate the use of additional authentication and authorization schemes for an existing infrastructure without having to update or modify multiple web services.
- Migration Towards Single-Sign-On
  - It is desirable for all web services associated with the CMCS to enjoy a single sign-on. We will continue to assess the costs and benefits of this feature as our web technologies mature.
- Support Additional Authentication/Authorization Schemes
  - It is desirable to extend the initial Implementation-Independent Security Model to allow configurations that support and/or require Public Key Infrastructure (PKI) and sophisticated Grid-based security. Such additional schemes will extend the reach of CMCS to communities that have additional security requirements.